DOCUMENT RESUME

ED 049 291 TM 000 478

AUTHOR Brown, Robert D.
TITLE Student Development in an Experimental College: Some Evaluation Strategies and Outcomes.
PUB DATE Feb 71
NOTE 11p.; Paper presented at the Annual Meeting of the American Educational Research Association, New York, New York, February 1971

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS Curriculum Evaluation, Educational Research, *Evaluation Techniques, *Experimental Colleges, Experimental Programs, Higher Education, Program Effectiveness, *Program Evaluation, *Research Problems, Student Attitudes, *Student Development
IDENTIFIERS Centennial Education Program, *University of Nebraska

ABSTRACT

Several problems that an evaluator of an experimental college has to face and resolve are considered. These include: (1) the purpose of the evaluation and the role of the evaluator, (2) the utility of a research design model, (3) the specificity of program goals before the evaluation is planned, and (4) the appropriateness of standardized or local assessment instruments. Evaluation problems of the Centennial Education Program at the University of Nebraska, which included 125 college freshmen and 50 upperclassmen, are discussed. Results of the evaluation revealed some clear differences in student development between those in the experimental college and those not in the program. Students in the program were found to have closer relationships to faculty and other students, read more non-required books, showed a greater preference for participatory academic activities, showed greater interest in reflective thought, an increased fondness for novelty, and were less practical and materialistic than their counterparts in the control group. No significant difference on any index of academic achievement or ability was noted. Random assignment of students was made in establishing the experimental and control groups. The use of a research model for assessment provided reliable indices of the effectiveness of the experimental program. (AE)

ED049291

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

Student Development in an Experimental College:

Some Evaluation Strategies and Outcomes*

Dr. Robert D. Brown
Coordinator of Research
on Student Development
University of Nebraska
Lincoln, Nebraska

INTRODUCTION

Evaluations of experimental colleges and innovative educational programs present unique challenges to persons taking on or being assigned the task of assessing outcomes. There is no readily available paradigm to serve as a guideline and though the thoughts of theoreticians provide clues for the justification of an evaluation strategy, they provide few notions that are helpful in the implementation of that strategy. As Theodore Newcomb has said "there is an almost desperate sense of need about better methods of evaluating (experimental colleges) in very much better ways than have ever yet been used." (Newcomb, 1970) Looking forward, one may wonder if these new innovative academic environments do not hold the potential to be the major advancement in higher education in the 1970s. They must, however, be better understood and properly evaluated, if their lessons are to have any semblance of permanence.

The purpose of this paper is to share some after-the-fact ruminations about the strategies employed in a particular evaluation enterprise and some of the outcomes. Most such reports primly summarize purposes, methods, and results leaving the impression the entire process was equally tidy.

TM 000 478

*Paper presented at the annual meeting of the American Educational Research Association, New York, N.Y., February 5, 1971

This summary will try to help the reader visualize the specter of dilemmas, crises, and quandries involved. The focus will be on several perplexities that an evaluator of an experimental college has to face and resolve. These include: 1) What is the purpose of the evaluation and the consequent role of the evaluator? 2) Should a research design model be employed? 3) How specific should the program goals be outlined before the evaluation is planned? and 4) Should standardized or local assessment devices be used? Before proceeding to these questions the nature of the experimental program will be briefly described.

Experimental Program

The Centennial Education Program initiated last year at the University of Nebraska offered a unique educational experience for 125 college freshmen and 50 upperclassmen. The program was established as an experiment in providing a living-learning environment which would foster the academic-intellectual and the personal-social growth of students. Features of the program included a residential setting with classes held within the same building as living quarters, a coed common lounge and recreation area, a core course of six hours offered on a pass-fail basis, and an emphasis on group and independent work on student chosen topics. Freshmen participants were randomly selected from a pool of eligible applicants.

STRATEGIES

Purpose and Role of Evaluation

The experimental college on the University of Nebraska campus was planned for several years before it came into being in the Fall of 1969. While educational psychologists were quite active in the planning stages of the college, which was later christened the "Centennial College," none

were formally associated with the college when it became a reality. The bulk of the faculty were from the humanities. Considering that humanities professors are often not in sympathy with the argot of educational psychology and that one of the planning committee members had decried the value of turning the college into an experimental lab, the initial task of the evaluator, an educational psychologist, became one of establishing credentials of worth and rapport.

There was no specific charge given to the evaluator. The University community -- administrators, faculty, and students -- were all interested in outcomes, but clues as to the nature of these outcomes or appropriate assessment techniques were unstated. There appeared to be an expectation that there would be an end of the year summative report, but the purpose of the evaluation was in a sense up to the evaluator to determine. Scriven (1967) has distinguished between formative evaluation, which attempts to provide on-going feedback to a program or project, and summative evalution, which might comment on the ultimate worth of the program. The focus of the efforts of the evaluator for this program was on the formative aspects. How might the faculty best be assisted to evaluate and change their behavior as the year went along and how might the outcomes of the first year be utilized to improve the college during the ensuing years? At the same time it seemed appropriate to take an omnibus approach and examine as many aspects of student development as possible.

The role options available to the evaluator ranged from being the hard-nosed researcher who would be demanding, exact, and impersonal to that of the participant-observer, who would be intimately involved with the experimental college and its community. One might well conjecture that the personal

style of the evaluator interacts with his philosophy and techniques and no doubt has an impact on the overall effectiveness, if not the worth of the evaluation. This question merits further empirical inquiry.

In this instance the evaluator chose a middle-of-the-road approach, not unlike his own personal style of interacting. It was surmised that a cold scientific manner would alienate the experimental college community, limit cooperation, and reduce the possibility that suggestions and comments would have any impact. On the other hand, full participation ran the risk of loss of objectivity and impressionistic conclusions would not be convincing to the rest of the University community. The evaluator did not become an active participant in the full sense, but did attend all staff meetings, all student town-meetings, and many college functions. There was participation, but comments were purposely limited to reactions, status reports, and very few suggestions. Seldom, if ever, did the evaluator espouse a particular resolution or course of action. During sessions with the entire staff or with individual staff members the evaluator assumed the same role as might the nondirective counselor, who listens and summarizes alternatives, but does not take a specific stance on an issue. This role might be labeled: consulting-evaluator.

Research Design

The relative merits of employing experimental design models in the evaluation of educational programs has become a debated issue. The arguments pro and con have direct relevance for evaluation of an experimental college. Cronbach (1963), for example, questions the value of having comparison groups, Stufflebeam (1970) believes that experimental design has limited usefulness depending upon the purpose of the evaluation, and Guba (1969)

questions the commonality of research and evaluation. Stanley (1969), on the other hand, is unhappy because research models have not been employed more often in evaluation projects and Scriven (1967) supports a control group evaluation model, if it is not the "total approach."

The strategy chosen for the evaluation of this project assumes that evaluation is something more than research. This does not mean that a research model or design is inappropriate as part of an evaluation, but it is just that -- a part, not the whole. There cannot be the extent of control of the many variables involved and the more control there is the stronger is the possibility that the college program will not truly reflect what its planners and participants intended. A strict research model eschews program changes during the experiment and the evaluator who limits himself to such a design is likely to overlook many valuable insights. However, for the evaluator to fail to attempt to control for as many invalidities as feasible would be comparable to a surgeon failing to wash his hands and instruments before attempting very risky surgery. Certainly such a rationalization is unethical, if not criminal.

These conclusions led the evaluator of the experimental college at Nebraska to include a research model as part of the evaluation scheme. The design chosen for this project was the separate-sample pretest-posttest control group design (Campbell & Stanley, 1966). This quasi-experimental design permitted the collection of pre-and post-test data on students without any one student taking the same instrument twice and yet making beginning and end of the year comparisons feasible. Because students were randomly assigned to the experimental college from a previously screened pool of students, it was also possible to make comparisons with students who had applied to the college and met the screening criteria,

but were not randomly selected and instead attended the regular University. Such a design made possible casual inferences about the impact of the total experimental college experience, although it was not possible to trace changes to specific aspects of the college program.

Specification of Goals

An interesting illustration of the complexities of on-the-site evaluation relates to how the goals of the program were derived, or at least stated. At one of the preparatory sessions two educational psychologists were invited to sit and discuss various strategies of assessing outcomes. During the discussion, as might be expected, the psychologists asked what some of the objectives of the college were to be. An almost visible shiver went up their spines when the reply came, "We don't know what our objectives are or what we are trying to do. You tell us what they should be." This response was not intended to be entirely in jest or to be sarcastic. At that point the program planners were clear that they did want to have an alternative educational program, but unclear as to just what it might try to accomplish that would be different from the traditional college curricula.

Pushing at this point for further delineation of objectives, much less asking that they would have to be in behaviorally expressed terms, would have only served to alienate the staff and would in fact have been absurd. It must be admitted that this kind of experience can be frustrating as well as challenging. A modest search and scrutiny of statements of educational philosophy written by the program planners and attendance to their casual remarks, however, made it apparent that there were implicit statements of goals, even if they were not stated in specific behavioral terminology.

As the year progressed, the faculty became more introspective about the question of goals, as well as methods. In mid-year they were able to describe in detail the characteristics of their ideal student when he completed the program.

Efforts to have faculty members specify their objectives in strict behavioral terms often serves to alienate them, besides taking a good deal of time. The evaluator must carefully consider these disadvantages. Scriven (1967) suggests that it is not improper for the evaluator to translate the implicit objectives into his own jargon. In this venture, as is undoubtedly true for other experimental colleges, the goals themselves were not static. Reliance solely on a research model would have limited the evaluation to assessment of only initial goals. Evaluators must be alert to the changing nature of objectives, as well as tactics.

Choice of Instruments and Assessment Devices

The selection of assessment devices for an evaluation project can evoke a minor dilemma. Devising local instruments involves time and talent. Unless a good deal of pretesting is done, the validity and reliability of such devices remains in doubt. On the other hand, while the use of standardized tools counteracts these problems, the instruments are seldom completely applicable. Because the tactics and goals of this experimental college were varied, comprehensive and to a degree unpredictable, it seemed appropriate to use an equally variable approach in assessing student development. Personal interviews, observations from participant observers, questionnaires, and self reports formed essential elements in the assessment process. Four major standardized questionnaires were also used: the Omnibus Personality Inventory, the College Student Questionnaire, the College

and University Environment Scales, and the Watson-Glaser Test of Critical Thinking Appraisal.

OUTCOMES

Usefulness of Evaluation Strategy

Before summarizing the student development outcomes, it seems appropriate in view of the preceeding discussion to comment on the effectiveness of the evaluation strategy employed; that is, use of the research model plus supplementary tactics. Perhaps the ultimate behavioral evidence of success is that the evaluator has been asked to serve in the same role again for the second year of the program. The evaluator has been more frequently asked to advise individual staff and students on projects and problems. As a result the evaluator is more than ever convinced that use of a research experimental design model which utilizes chiefly tests and questionnaires is too narrow. Observation, interviews, and some degree of interaction with the participants are vital.

The novelty effect is difficult to control without either a placebo treatment or another experimental college of a different variety. Hopefully, someday comparisons will be possible between distinct experimental programs within and between colleges. Until then such programs must be compared to what they could be and mini-evaluations of specific dimensions should be encouraged. Next year the evaluator is considering some use of matrix sampling (Husek, 1968) and educating the staff to the advantages of the systems approach suggested by Axelrod et. al. (1969).

The summative end of the year report consisted of a summary of data, observations, conclusions and suggestions. The report was well-received, even though not all of its conclusions were universally accepted. Since then the staff of the college has devoted considerable time and effort to following up on the suggestions.

Experimental Outcomes

The results of this study revealed some clear differences in student development between those in the experimental college and those not in the program, provided some significant distinctions between the most and least successful students, and offered some suggestions for program improvement.

Living-Learning: The effectiveness of the living-learning environment was supported by questionnaire and interview data indicating that the experimental college students had a closer relationship with their faculty members than did regular University students. They also had a broader, though somewhat less intense, relationship with other students and tended to be involved in more intellectual, informal bull-sessions. Interviews also revealed that program students sought even more informal contacts with faculty.

Though there were no differences between the experimental college students and the control group in their ability to think criticially, there were significant differences in their academic styles and intellectual orientations. Program students read more books related to their courses that were not required, were more familiar with the library, more took notes when they read and more showed a greater preference for participatory academic activities, such as independent work, original research and class discussions. They also showed greater interest in reflective thought and an increased fondness for novelty.

Personal-Social Values: The experimental students became more liberal and more independent of authority, but no more so than the regular University students. They did become significantly less practical and materialistic. A significantly greater percentage of the experimental group felt their greatest problem was their quest for identity and more were uncertain about their future vocational goals. Interviews suggested that the program

students experienced change but had little direction, while control students denied change, but felt a greater awareness of others' values.

Student Characteristics: Students ranked as more successful by faculty members were more reflective and intellectual and at the same time more socially extroverted than were students portrayed as maturing less. There were no differences, however, on any indices of academic achievement or ability. Faculty impressions of intellectual growth correlated significantly with student grades, intellectual orientation, social extroversion and altruism, but not with ability indices. Assessments of personal-social growth followed the same pattern. Students electing a pass-fail grading option were found to be less intellectually oriented and fewer were risk-takers.

IMPLICATIONS

The use of a research model for assessment provided reliable indices of the effectiveness of the experimental program. This was possible because of the random assignment of the students to the experimental and control groups. At the same time the supplementary evaluation tactics proved invaluable in making the test data richer and providing insights about specific ways in which the program could be improved. Such a two-pronged approach for evaluating experimental programs appears to be more useful than the strictly research model alone.

References

Axelrod, Joseph, Freedman, Mervin B., Hatch, Winslow R., Katz, Joseph and Sanford, Nevitt Search for Relevance, Jossey-Bass, San Francisco, 1969.

Campbell, Donald & Stanley, Julian Experimental and Quasi-Experimental Design, Rand McNally & Co., Chicago, 1966.

Cronbach, Lee "Evaluation for Course Improvement." Teachers' College Record, 64, No. 8, May, 1963, 672-683.

Guba, Egon C. Significant differences. Educational Researcher, 1969, 20, (No. 3, 4-5).

Husek, T. R. & Sirontnik, Ken "Matrix Sampling" Evaluation Comment, December, 1968.

Newcomb, Theodore Personal correspondence, April, 1970.

Scriven, Michael "The Methodology of Evaluation," in Perspectives of Curriculum Evaluation, American Educational Research Association Monograph Series, Rand McNally & Co., Chicago, 1967.

Stanley, Julian "Reactions," Educational Researcher, 1969, 20, No. 5, 8-9.

Stufflebeam, Daniel L. "The Use of Experimental Design in Educational Evaluation," paper read at AERA, Minneapolis, March, 1970.